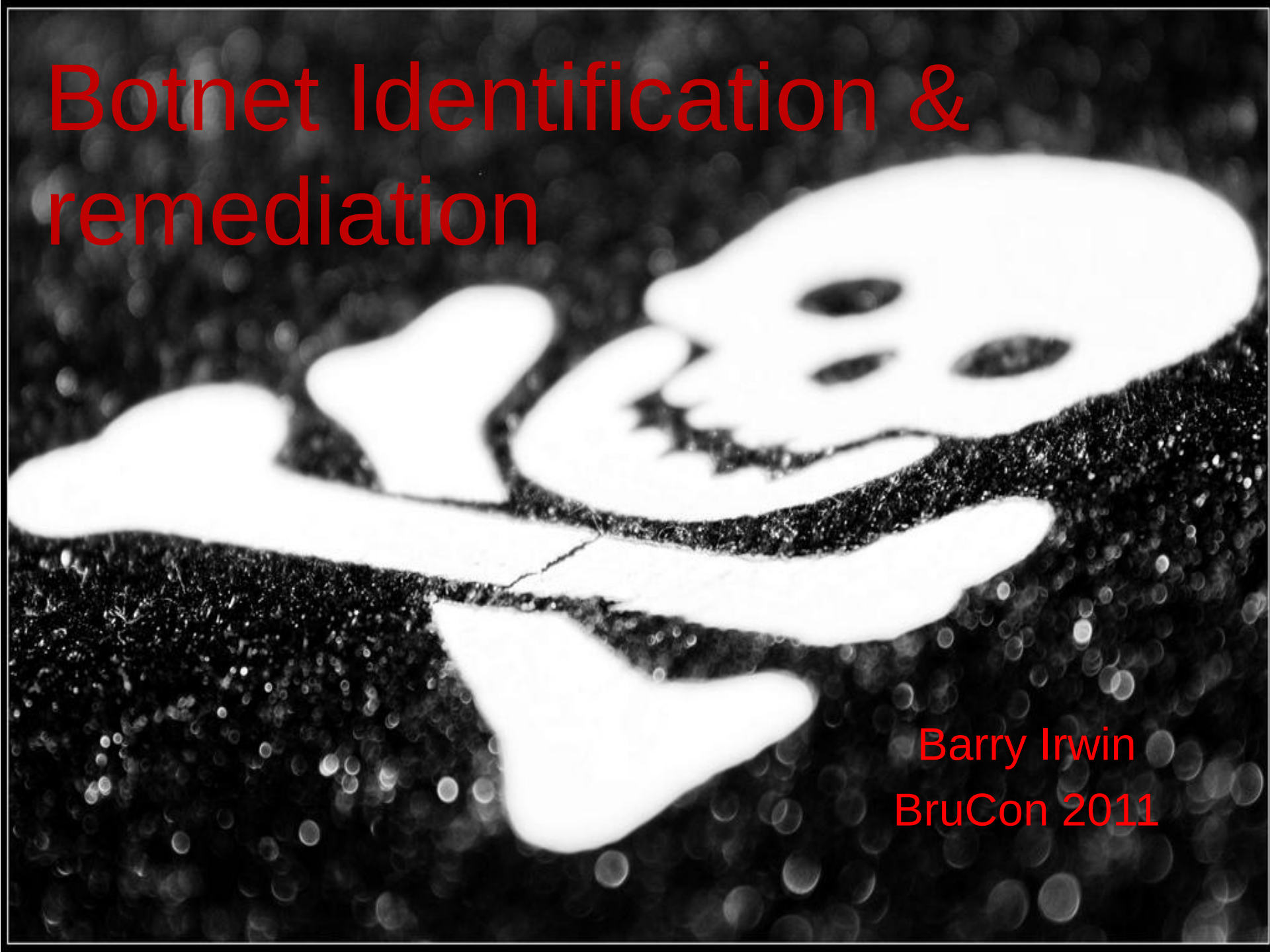
Botnet Identification & remediation



Barry Irwin

BruCon 2011

About

RHODES UNIVERSITY
Where leaders learn

SNRG
Security and Networks Research Group

Overview

- Why ?
- How ?
  - DNS
  - Lexical
  - Visualization
- What ?
- When ?

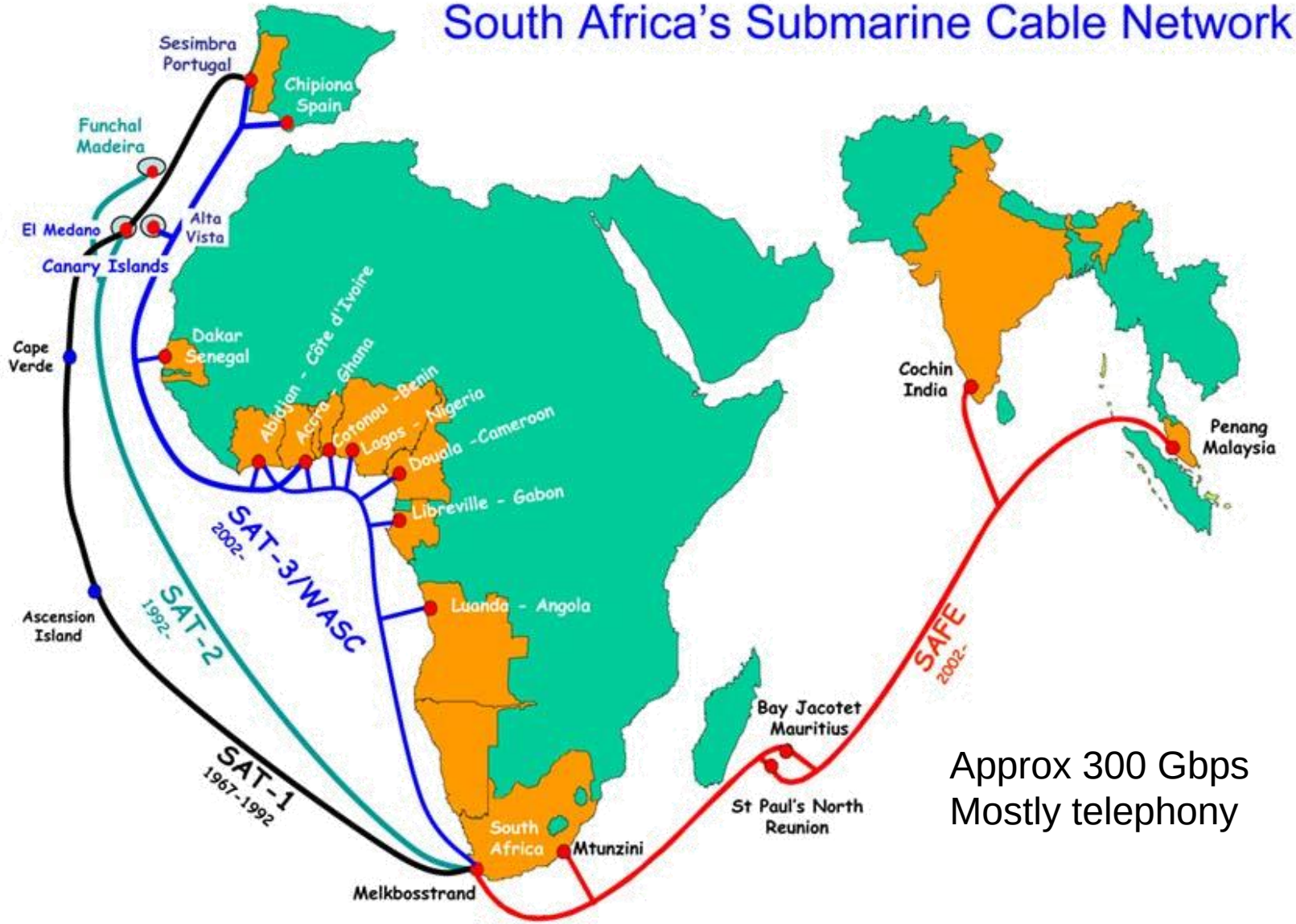
Background

Why should we bother looking at this ?

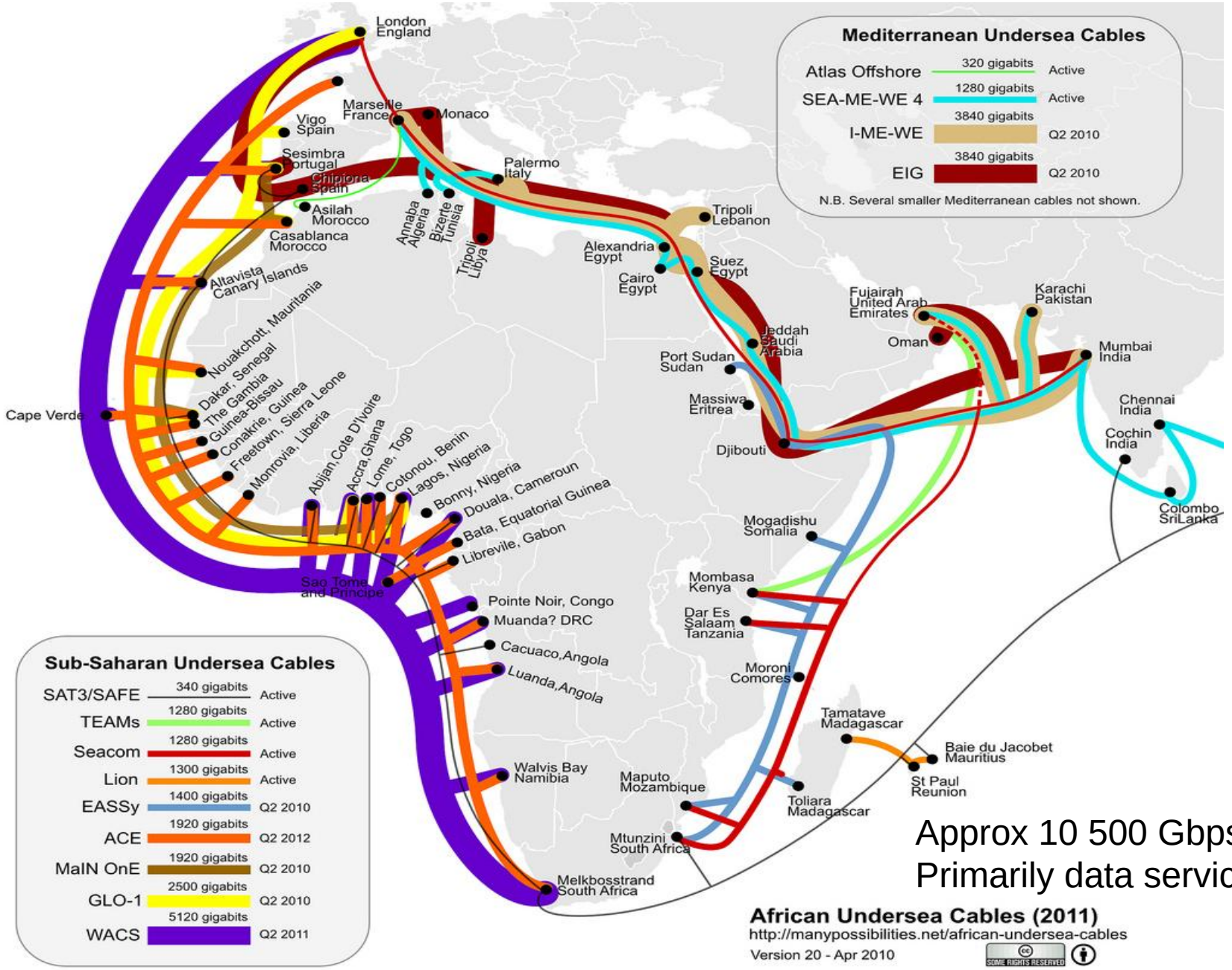
The Scene (and how we got started)

- African Countries are facing a new challenge (and big it is!)
- Potential for massive uptake & growth in Internet connectivity
- We are facing increased risk
- Are we actually equipped and resourced ?
- This is NOT an African problem!

Pre 2007



Approx 300 Gbps
Mostly telephony



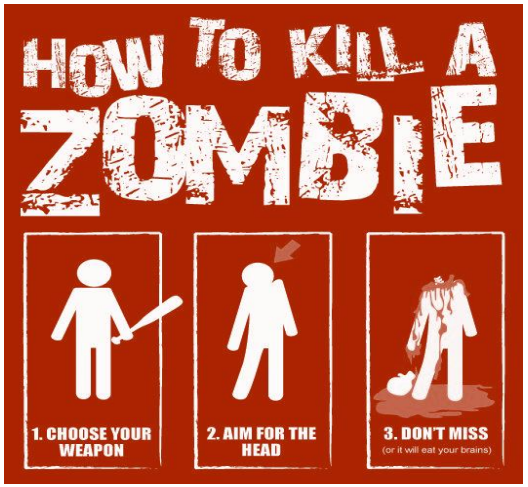
Approx 10 500 Gbps
Primarily data service

Botnet Evolution

- Modern botnet trends have become increasingly sophisticated both in terms of the techniques used to avoid detection on compromised endpoints, but also in their varied communication channels.
- Traditionally, IRC used as Command & Control (C2)
- New malware uses IP and domain fast-fluxing to avoid detection and increase resilience.
  - DNS tunnels, DNS C2 , HTTP, P2P
- These techniques largely bypass traditional network security detection and mitigation approaches

Mitigation

- Traditional approaches - resource intensive
- Difficulty in automating
- Difficult to ‘close the loop’ & apply remediation



- Not so Easy for Digital Zombies

Why bother ?

- Isnt this someone else's problem ?
- Its $vendor’s fault
- My AV is almighty?
- It wont happen to me!

- If you don’t care about the ‘small’ things…..

Defense

Words of wisdom

It is essential to seek out enemy agents who have come to conduct espionage against you and to bribe them to serve you.

Hold out baits to entice the enemy. Feign!

-- Sun Tzu

The DNS approach

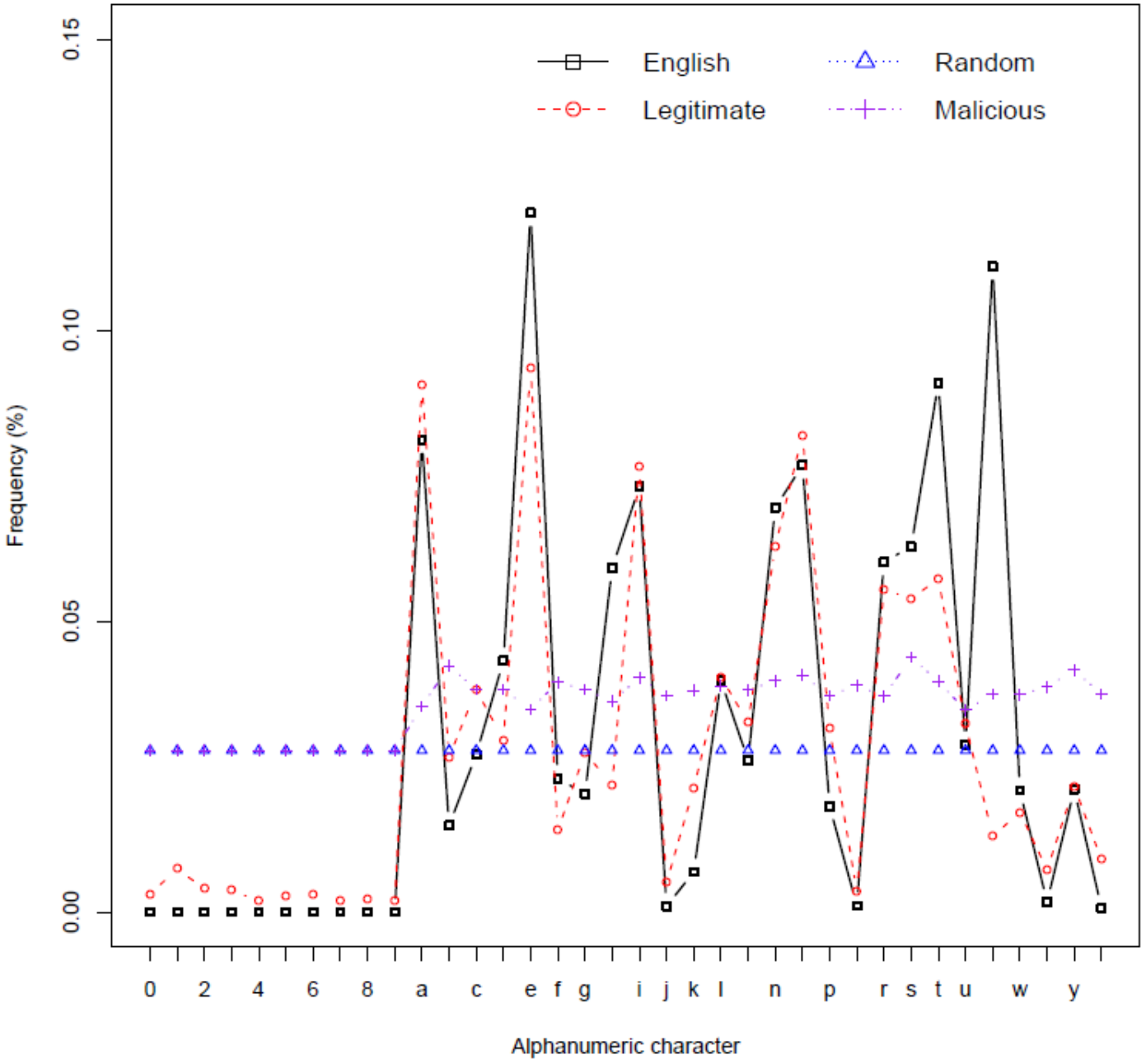
- DNS is tightly integrated into the TCP/IP protocol stack
- Commonly being used as a means of pointing to C2 servers
- Provides a fixed → volatile mapping
- Essential, but commonly ignored protocol
- Completely passive collection of DNS query data can be done on a network
  - Tap
  - DNS Servers

Passive lexical analysis

- Actual Domain names can be analysed and scored
- Domain names provide a readable and easy to remember mapping between a domain and its IP data
- Typically consists of English words or a combination of English words.
- Malicious domains, which are algorithmically generated commonly do not contain English words or letter combinations usually seen in legitimate domains
  - Conficker (xllnm.com.do)
  - Kraken (ygcoqgmmb.yi.org)

Making decisions

- Possible to calculate letter frequencies as they occur in DNS
  - legitimate,
  - randomly generated
  - algorithmically generated domain names.
- Statistical methods are applied to these frequency distributions
- Enables allowing for accurate classification of domain names as legitimate or malicious, based solely on alphanumeric character distribution


English
Legitimate
Random
Malicious
Frequency (%)
0.00
0.05
0.10
0.15
0 2 4 6 8 a c e f g i j k l n p r s t u w y
Alphanumeric character

Interactive Queries

- Domains can be interrogated often with a use of a single query
- Extract pertinent data
  - A Records
  - NS Records
  - Network Ranges
  - Unique ASNs
  - TTL
- This is found to have a strong correlation to the domain ‘classes’

DNS attributes

	Standard DNS	CDN	Fast-Flux
A Records	4	4	4
NS Records	2	2	2
Network Ranges	1	1	3
Unique ASNs	1	1	2
TTL	≥ 1800	< 1800	≤ 600

- Fast-flux domains share many common characteristics with CDNs
- Sample data shows that fast-flux domains have the shortest average TTL
- Furthermore, the domain hosts are spread across multiple, widely dispersed IP ranges.
- The domain hosts are associated ASNs.
- CDNs display similar characteristics, but on average are associated with fewer ASNs and resolve to less widely dispersed IP ranges.

Sample Fastflux Query

championogogo.ru				
IP Address	Net block	ASN	Country	TTL
60.13.74.23	60.13.64.0/18	4837	CN	300
62.42.100.212	62.42.0.0/16	6739	ES	300
148.217.94.55	148.217.0.0/16	6503	MX	300
212.69.189.125	212.69.160.0/19	8218	DE	300
217.217.199.129	217.216.0.0/15	6739	ES	300

Other attributes to consider - Whois

- The nature of fast-flux domains dictates that they are associated with recently registered domains.
- Be obtained through a whois query.

  day old bread – already used for spam detection
- Information returned by
- Whois query includes
  - registration date,
  - the registration authority
  - (potentially) the country the domain was registered in
- These are all potential data points for classification

Making it smart!

- DNS data elements can be fed into a classification system, requiring just a single DNS query
  - C5.0 Decision tree classifier
  - Naive Bayesian classifier (trained)
- #include <very_scary_maths.h>

$$ln\frac{P(F \mid D)}{P(\neg F \mid D)} = ln\frac{P(F)}{P(\neg F)} + \sum_i ln\frac{P(t_i \mid F)}{P(t_i \mid \neg F)}$$

$$\sigma(P,Q) = \frac{1}{2}\sum_i \mid P(x) - Q(x) \mid$$

Initial Results - Bayes

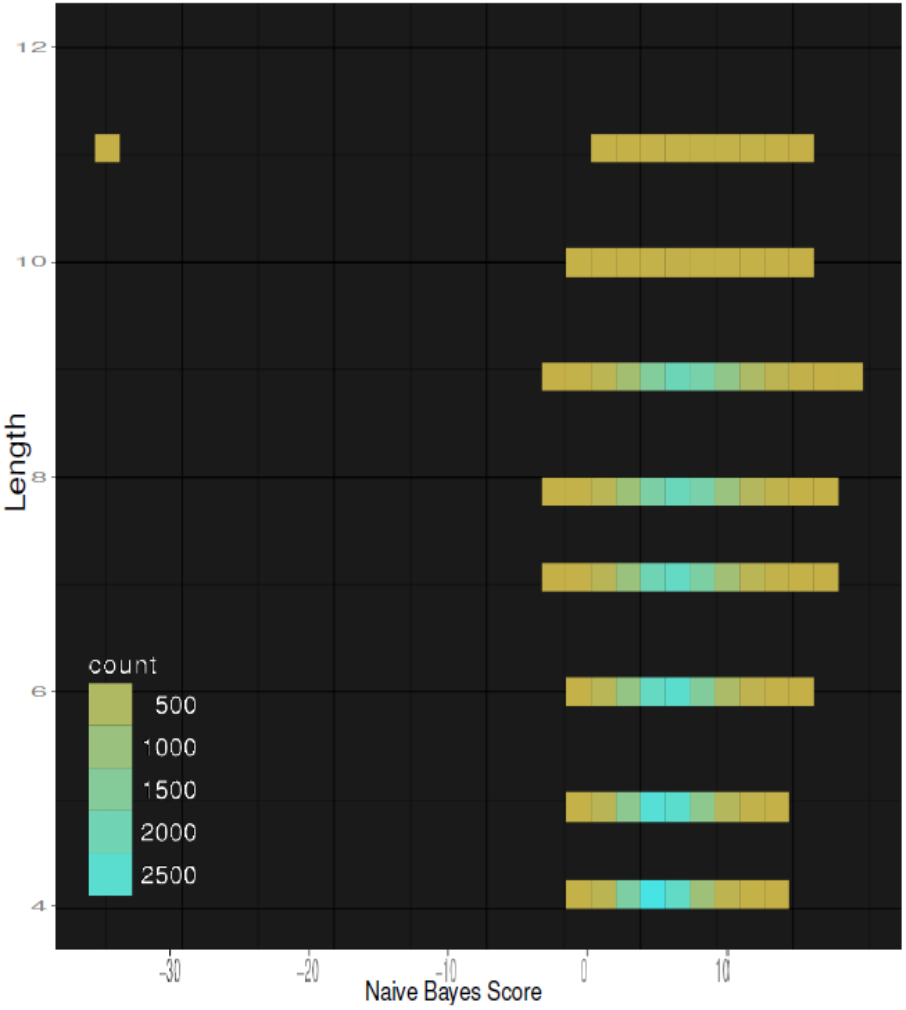
- Training Data

	A Records	NS Records	Number IP ranges	Number ASNs	TTL
Fast-Flux	2.090032	3.916399	2.180064	3.70418	594.9968
Legitimate	1.730769	3.87574	0.1538462	1.094675	14885.42

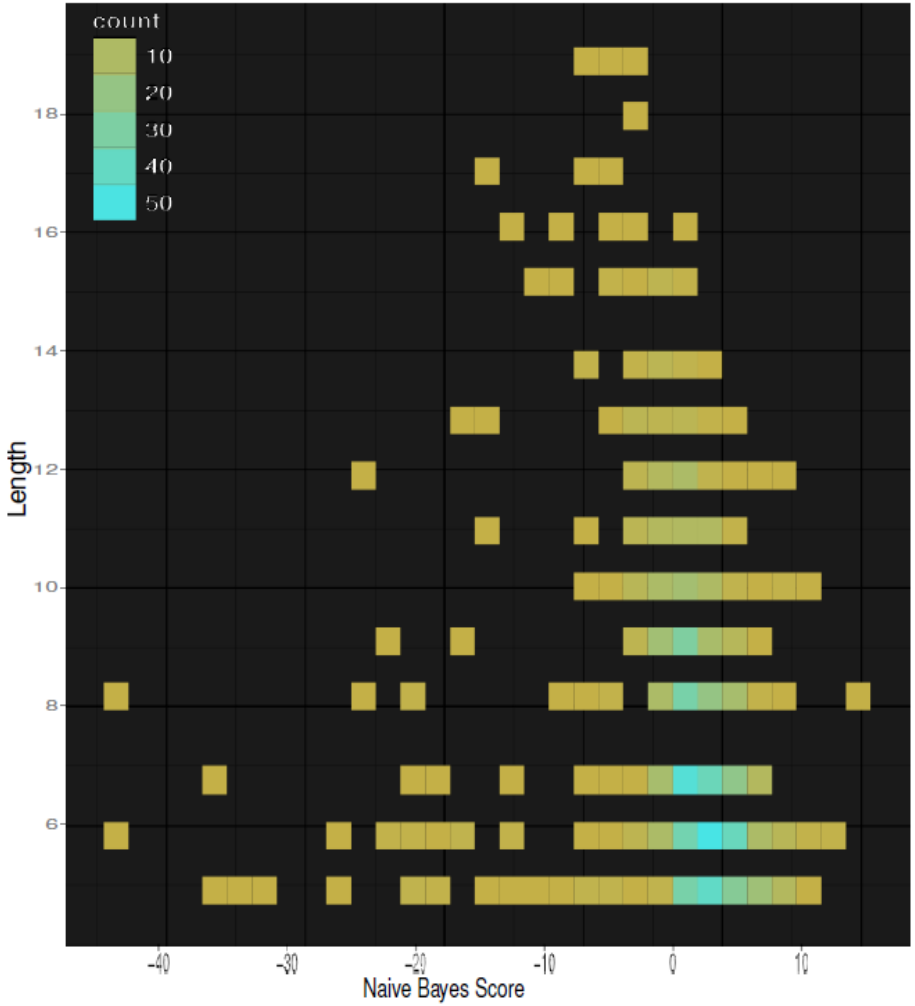
- Test Results

Domain	Safe Score	Malicious Score	Classification
gingerbucksea.com	0.005304578	0.3550235	Fast-flux
pearlrumor.ru	3.059976e-14	7.490562e-13	Fast-flux
wordpress.com	1.536894e-08	4.250896e-10	Legitimate
champiogogo.ru	3.395984e-09	1.723838e-06	Fast-flux
yahoo.com	1.940412e-15	1.509179e-69	Legitimate

Bayes Scoring



(a) Malicious domains



(b) Legitimate domains

Bayes Scoring Results

Domain name	Output	Classification	Correct
Facebook.com	-1.06400	Legitimate	Yes
allrecipes.com	-4.25654	Legitimate	Yes
twitter.com	-2.39181	Legitimate	Yes
buzzle.com	2.47540	Malicious	No
nhk.or.jp	0.64375	Malicious	No
bbhkxkjh.com.fj	6.61512	Malicious	Yes
pveufjtm.com.bo	3.25285	Malicious	Yes
rrxwigqj.am	5.24226	Malicious	Yes
ljtkrinq.com.tt	2.75078	Malicious	Yes

URL Based Heurisitcs

- The second component applies a lightweight mathematical classification to observed URLs contained in network traffic.
- This can either be via a network tap, or integrated into a proxy server, or browser plugin
- The methods used are able to identify malicious urls with a high degree of accuracy, while maintaining a low false positive rate.
- This lightweight solution can be further supported by active queries

Lightweight URL classification

- Recent trends indicate motivation behind malicious websites has moved towards financial gain.
- Primarily done through the use of phishing sites as well as spam sites that attempt to sell fake goods such as pharmaceuticals
- Increasingly through the use of `drive-by-downloads`
- The main outcomes of malicious content:
  - Phishing
  - Fraudulent advertising
  - Malicious downloads
- Key advantages of being able to identify these URLs
  - Blacklisting
  - Incident analysis
- For an telco provider, speed is a crucial factor when implementing an on-the-fly countermeasure of this kind

Obfuscation

- Type I
  - This type of obfuscation refers to cases where the hostname is replaced with an IP address and/or a port number is used.
- Type II
  - The hostname in the URL has a domain name that appears to be legitimate but usually contains a redirect to another host.
- Type III
  - The host name is obfuscated, but in this form a large string of other valid domains is appended to it.
- Type IV
  - The domain name is misspelled or no domain name is given in the URL.

Current Countermeasures

- Blacklists
  - Phishtank
  - OpenDNS
  - Day old Bread
  - Google Safebrowsing
- These do not necessarily respond fast enough.
- Automted Classification
  - Host based
  - Lexical
  - Full Featured

URL Classification

- Host-based features
  - These use external data source
    - WHOIS
    - Team Cymru.
    - ASN
    - Blacklists
    - Heuristic tagging (as we just discussed)
    - Remote OS
    - Remote portlists
    - **Any** other external source
  - The negative aspect of using these external features for classification is that they may incur significant additional latency.

URL Classification

- Lexical features
  - Process the actual text of a URL. (excludes host)
  - Malicious URLs often ”look” different to experts when compared to benign ones
    - Phishing
    - Malware uses ‘randomish’ url paths
    - http C2 nodes sometimes have predictable components
  - Features that are commonly used in this type of classification include numerical information regarding lengths of features, numbers of delimiters and path structure
- No Semantic information

Implementation

- The three classification algorithms to be implemented
  - Online Perceptron
  - Confidence Weighted (CW)
  - Adaptive Regularization of Weights (AROW)
- All are single layer neural networks
  - One input layer and one output layer using linear combiner)function
  - #include <more_headache_maths.h>

$$(\mu_{t+1}, \Sigma_{t+1}) = \arg\min_{\mu\Sigma} D_{KL}(\mathcal{N}(\mu, \Sigma) \| \mathcal{N}(\mu_t, \Sigma_t)),$$
$$\text{s.t.} \Pr_{w \sim \mathcal{N}(\mu,\Sigma)}[y_t(w \cdot x_t)] \geq \eta$$

$$(\mu_{t+1}, \Sigma_{t+1}) = \arg\min_{\mu\Sigma} D_{KL}(\mathcal{N}(\mu, \Sigma) \| \mathcal{N}(\mu_t, \Sigma_t))$$
$$+\lambda_1 l_h{}^2(y_t, \mu \cdot x_t) + \lambda x_t^T \Sigma x_t,$$
$$\text{s.t.} \Pr_{w \sim \mathcal{N}(\mu,\Sigma)}[y_t(w \cdot x_t)] \geq \eta$$

Initial Perceptron Results

- Online Perceptron is the only method currently complete
- Input Data
  - 5234 malicious URLs from Phishtank
  - 5408 benign URLs from Open Directory
  - Split into two sets for training and testing
- Perceptron has been implemented in C#
  - Currently converges on training data with a >99% accuracy.
  - The false positive rate is 0.1%
  - False negative rate is 0.4%.
- Classification times for test run <10 s for >4000 urls
- Further testing is required using real world data to establish a true error rate for this classification method.
- Obtaining Training Data is problematic

Initial Perceptron Results

- Online Perceptron is the only method currently complete
- Input Data
  - 5234 malicious URLs from PhishBank
  - 5408 benign URLs from Open Directory
  - Split into two sets for training and testing
- Perceptron has been implemented in C#
  - Currently converges on training data with a >99% accuracy.
  - The false positive rate is 0.1%
  - False negative rate is 0.4%.
- Classification time for test run <10 s for >4000 urls
- Further testing is required using real world data to establish a true error rate for this classification method.
- Obtaining Training Data is problematic

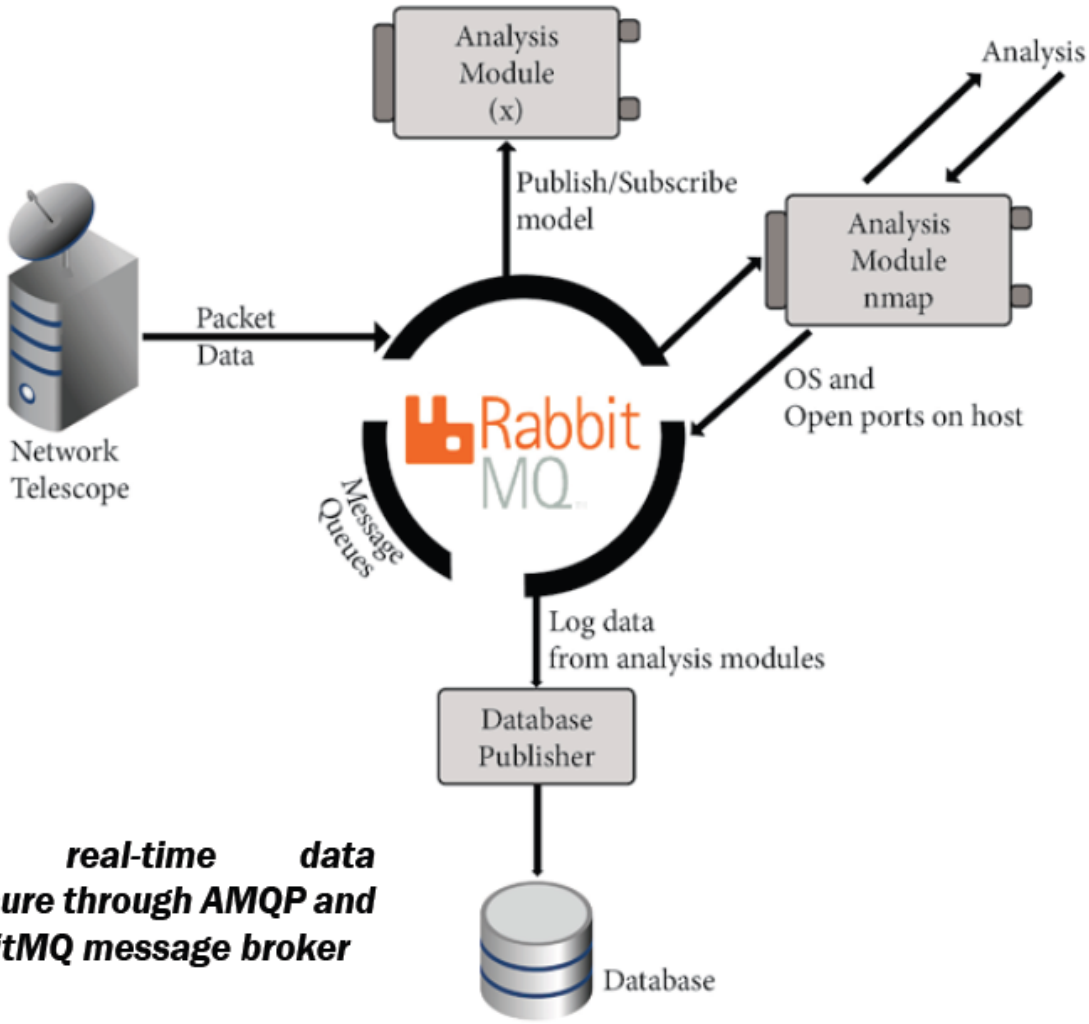
This almost looks TOO good!

Remediation

Monitoring

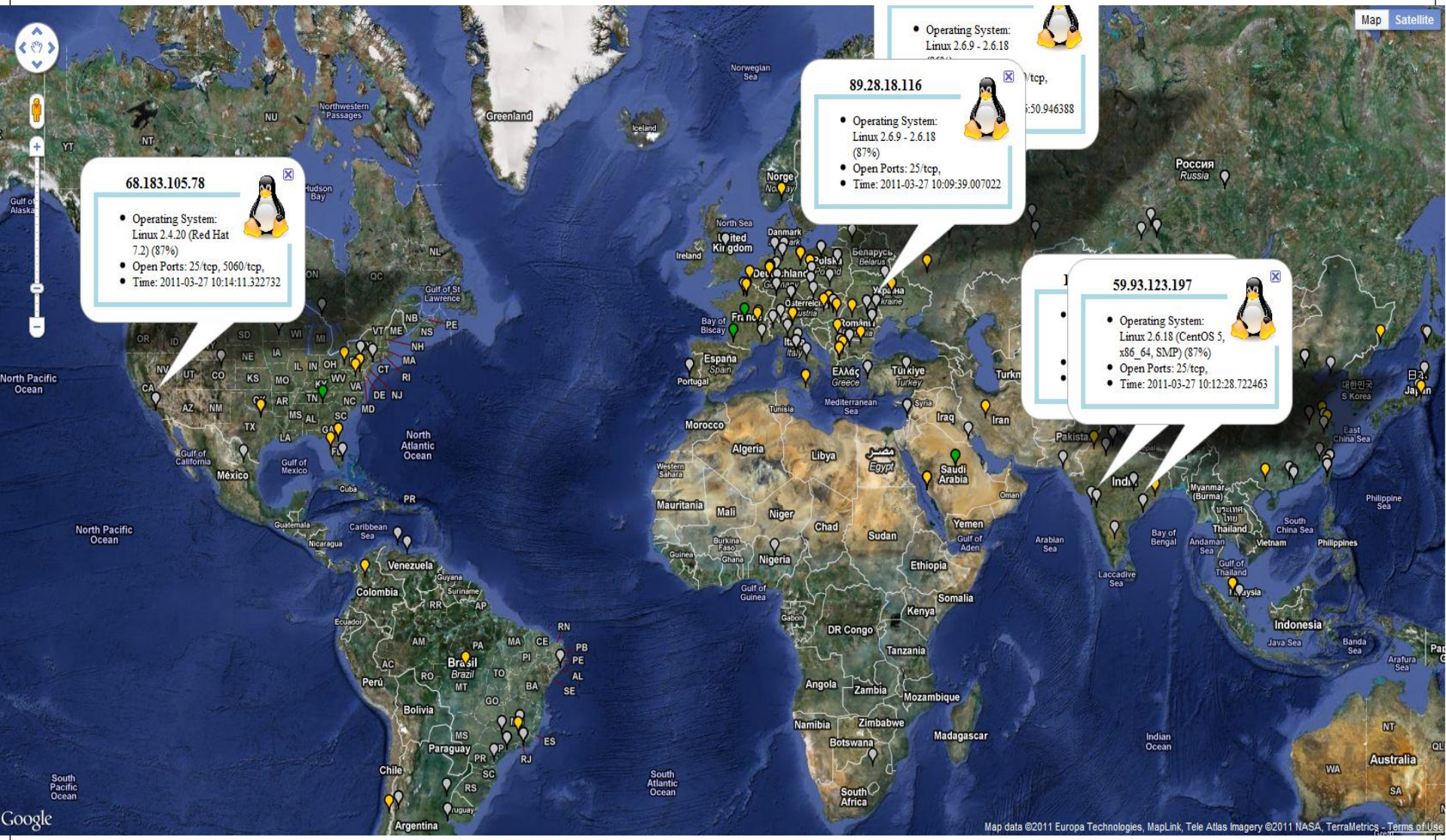
- Key element in remediation is being able to manage the data
- Integrate the components discussed in a web visualization framework (Google maps)
- The provides:
  - Overview
  - Geographic information
  - Further Details
- Re-purposed much of the work & expertise done on Network Telescope Research in the last 7 years

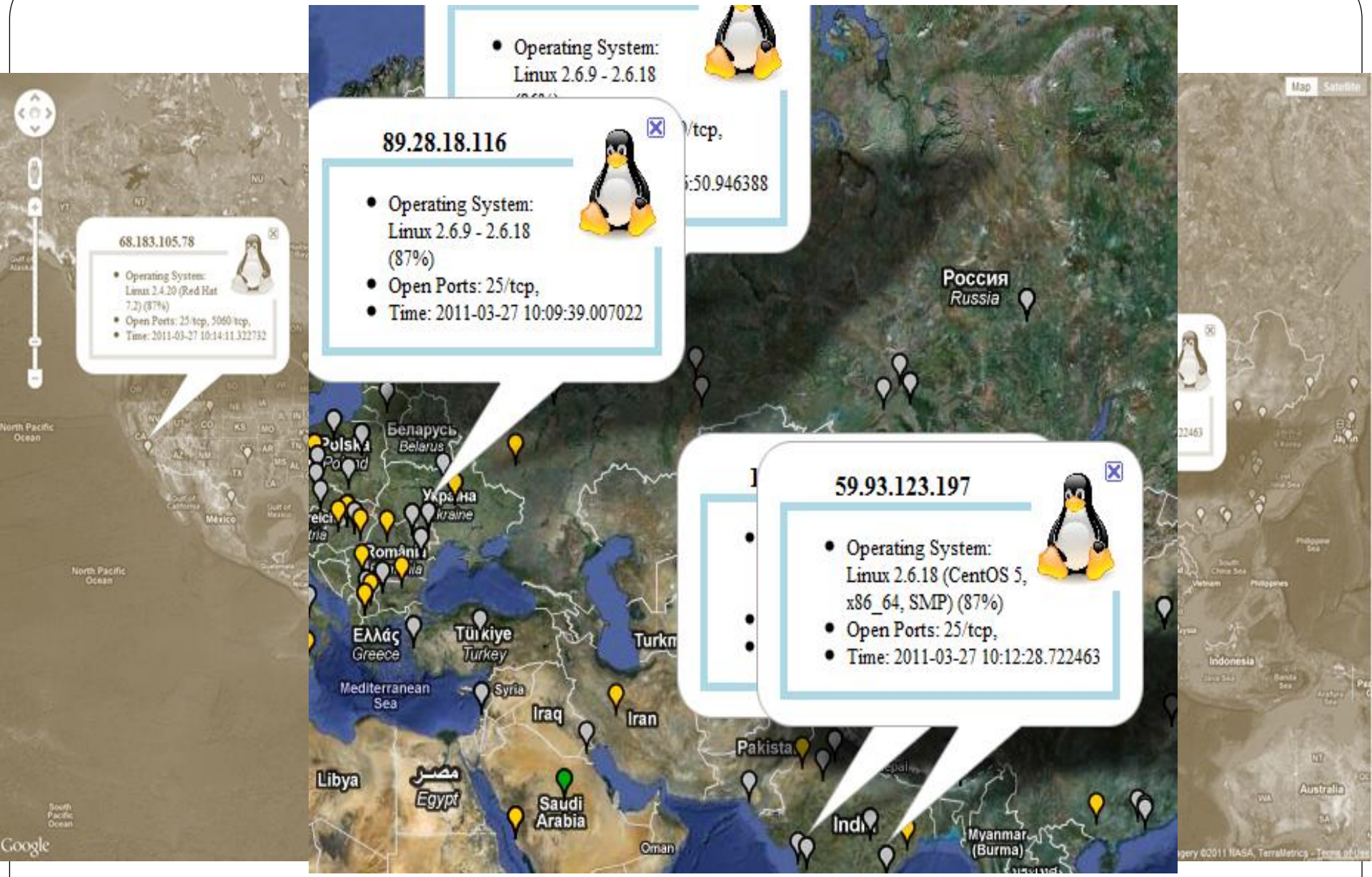
Architecture



Near real-time data exposure through AMQP and RabbitMQ message broker

Visualisation


68.183.105.78
Operating System: Linux 2.4.20 (Red Hat 7.2) (87%)
Open Ports: 25/tcp, 5060/tcp,
Time: 2011-03-27 10:14:11.322732
89.28.18.116
Operating System: Linux 2.6.9 - 2.6.18 (87%)
Open Ports: 25/tcp,
Time: 2011-03-27 10:09:39.007022
Operating System: Linux 2.6.9 - 2.6.18
59.93.123.197
Operating System: Linux 2.6.18 (CentOS 5, x86_64, SMP) (87%)
Open Ports: 25/tcp,
Time: 2011-03-27 10:12:28.722463
Map
Satellite
Google
Map data ©2011 Europa Technologies, MapLink, Tele Atlas Imagery ©2011 NASA, TerraMetrics - Terms of Use


68.183.105.78
Operating System: Linux 2.4.20 (Red Hat 7.2) (87%)
Open Ports: 25/tcp, 5060/tcp,
Time: 2011-03-27 10:14:11.322732
89.28.18.116
Operating System: Linux 2.6.9 - 2.6.18 (87%)
Open Ports: 25/tcp,
Time: 2011-03-27 10:09:39.007022
59.93.123.197
Operating System: Linux 2.6.18 (CentOS 5, x86_64, SMP) (87%)
Open Ports: 25/tcp,
Time: 2011-03-27 10:12:28.722463
Россия
Russia
Беларусь
Belarus
Ελλάς
Greece
Türkiye
Turkey
Mediterranean Sea
Syria
Iraq
Iran
Libya
مصر
Egypt
Saudi Arabia
Oman
Pakistan
Myanmar (Burma)
Indonesia
Australia
North Pacific Ocean
Google

Operational Use

- Can be used to track C2 servers
  - Further probes on these hosts
- Client/Bot tracking
  - Particularly useful to large/Regional ISP’s
  - Notification/Remediation Status tracking
- Live/Killed/Dead tracking

Remediation is HARD

- Identifying C2 nodes and local clients is the easy part
- Getting these cleaned up requires significantly more effort and resources
- Cutting the heads off a Hydra
- Endpoint apathy
- Relies on
  - CIRTS, Operators, Law Enforcement, Legislators

Bringing it together

When do we start ?

“ The world is a dangerous place, not because of those who do evil, but because of those who look on and do nothing.”

- Albert Einstein

Our solution

- Still very much work in progress
- Some very promising results
- Code Releases to come
- Browser plugins for Chrome/Firefox
- These techniques need to be forged into operational tools and process

Cleaning up

- Call to action or maybe to arms ?
- Multiple and diverse players required
  - Government
  - Internet Industry
  - Public-Private Partnerships
  - Grass roots education
- We have already seen the results of inaction.
  - Not necessarily malicious
  - rfc-ignorant.org (SMTP)
- Reputation and Trust

Where to start

Take the mote out of your own eye

Client Perspective

- Most people are only concerned where their money is affected
- Yet issues like phishing are still a major problem
- People Love the dancing bears.....
- Warnings can be provided, but as long as its on an endpoint system it can be ignored/cancelled/uninstalled
- User Education is a HUGE challenge
  - Three strikes for malware ?

Operator Perspective

- Reputation is critical for reliable network operations
- Methods describes, can provide a low resources means of flagging potently malicious activity with a fairly high accuracy rate.
- Internal Remediation/mitigation ?
- Dealing with Customers
  - Follow the model in Germany ?

InfoSec Research Perspective

- New tools allow one to look at data in a different light
- With understanding come insight
- DNS is a very unloved protocol ito monitoring
- Botnets are not about to go away, but will probably get more numerous

Start Small and grow

- Lock down Botnets/Malware
- Deal with phishing
- Spam
- Self-Regulation within industry groups
- Operator - ‘name and shame’ ?
- Consumer Education
- How do you/your organisation play nicely in the Internet Sandpit?

- Inaction is likely to have consequences

How do you/your organisation play nicely in the Internet Sandpit?

Inaction is likely to have consequences

Thanks

Etienne Stalmans
Samuel Hunter
Shaun Egan

More Detail (Homework)

- ***A Framework for DNS Based Detection and Mitigation of Malware Infections on a Network***. E Stalmans & B Irwin. 9th Information Security South Africa (ISSA) Conference, August 2011
- ***An Evaluation of Lightweight Classification Methods for Identifying Malicious URLs.*** S Egan & B Irwin. 9th Information Security South Africa (ISSA) Conference, August 2011
- ***Tartarus: A honeypot based malware tracking and mitigation framework.*** S Hunter & B Irwin. 9th Information Security South Africa (ISSA) Conference, August 2011
- ***Near Real-time Aggregation and Visualisation of Hostile Network Traffic.*** S Hunter & B Irwin. 14th Southern African Telecommunications and Network Applications Conference (SATNAC), September 2011
- ***High Speed Lexical Classification of Malicious URLs***. S Egan & B Irwin. 14th Southern African Telecommunications and Network Applications Conference (SATNAC), September 2011
- ***A Framework for DNS Based Detection of Botnets at the ISP level***. E Stalmans & B Irwin. 14th Southern African Telecommunications and Network Applications Conference (SATNAC), September 2011

Questions

Barry Irwin

b.irwin@ru.ac.za

@barryirwin